

MULTI-PROCESSOR SYSTEM RECOVERY
USING THERMTRIP SIGNAL

Inventor: Martin McAfee
1909 Patriot Drive
Lago Vista, Texas 78645

Kevin L. Miller
10710 Bull Ridge Drive
Austin, Texas 78759

Robert L. Nance
219 Serenada Drive
Georgetown, Texas 78628

Robert S. Tung
1607 Lobo Mountain Lane
Round Rock, Texas 78664

Assignee: DELL PRODUCTS L.P.
One Dell Way
Round Rock, Texas 78682-2244

BAKER BOTTS L.L.P.
One Shell Plaza
910 Louisiana
Houston, Texas 77002-4995

Attorney's Docket: 016295.1342
(DC-04829)

ATTORNEY'S DOCKET
016295.1342
(DC-04829)

PATENT APPLICATION

1

MULTI-PROCESSOR SYSTEM RECOVERY
USING THERMTRIP SIGNAL

TECHNICAL FIELD

This invention relates to processing systems, and more particularly to heat monitoring for processing systems with multiple processing nodes.

5

BACKGROUND

Most of today's processors incorporate a temperature sensor used for thermal monitoring. Often, the thermal monitor is integrated into the processor silicon. It includes a temperature sensing circuit and means for generating a signal (PROCHOT) that indicates that the processor has reached a maximum safe operating temperature. The processor may also include control circuitry that can automatically reduce processor speed and thereby reduce power consumption while the processor temperature is high.

In addition to the PROCHOT signal, or perhaps, alternatively, processors may also include an on-die diode that monitors the die temperature (junction temperature). If the temperature rises above a predetermined threshold, the processor shuts down. More specifically, when the junction temperature rises above a certain temperature (i.e., 135 °C for the Pentium III processor), the processor stops executing all instructions. The processor signals this condition to the rest of the system with a THERMTRIP (thermal trip) signal. The processor will remain stopped until a reset signal goes active via a restart or reset switch.

SUMMARY

In accordance with teachings of the present disclosure, a system and method are described for responding to a thermal trip signal from a processor of a multi-node system. A temperature monitor is connected to receive a thermal trip signal from each processor. The temperature module is also connected to deliver an enable signal to a voltage control module associated with each node. The voltage control module is operable to deliver voltage to all processors of the node when the enable signal is on and to shut off power to all processors of the node when the enable signal is off.

If a processor becomes overheated and asserts a thermal trip signal, the temperature monitor receives the thermal trip signal, turns off the enable signal to voltage control module of the node containing the overheated processor, and delivers a system power signal to the chipset of the computing system. The system is then reset, such that all nodes other than the node containing the overheated processor regain power.

An advantage of the invention is that after a thermal trip signal from any one processor, the system may become operational even if the overheated processor remains overheated or otherwise inoperable. After a reset, the node with the overheated processor remains shut down as a result of the thermal trip signal, but the remaining nodes are in operation. The overall result is increased availability of the system, which is very important for systems such as high end servers.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete understanding of the present embodiments and advantages thereof may be acquired by referring to the following description taken in

5 conjunction with the accompanying drawings, in which like reference numbers indicate like features, and wherein:

FIGURE 1 illustrates a multiple processor system having a temperature monitor in accordance with the invention.

10 FIGURE 2 further illustrates the temperature monitor of FIGURE 1.

FIGURE 3 illustrates a method of responding to a THERMTRIP signal in accordance with the invention.

DETAILED DESCRIPTION

FIGURE 1 illustrates a server system 100 having two nodes 101 (Node A and Node B) and a temperature monitor 103 in accordance with the invention. By "server system" is meant a computing system on a network that manages network resources.

Although the following description is in terms of monitoring processors of a server system, the same concepts could be applied to any "information handling system" having multiple processing nodes, each node having one or more processors. For purposes of this disclosure, an information handling system may include any instrumentality or aggregate of instrumentalities operable to compute, classify, process, transmit, receive, retrieve, originate, switch, store, display, manifest, detect, record, reproduce, handle, or utilize any form of information, intelligence, or data for business, scientific, control, or other purposes. For example, an information handling system may be a personal computer, a network storage device, or any other suitable device and may vary in size, shape, performance, functionality, and price. The information handling system may include random access memory (RAM), one or more processing resources such as a central processing unit (CPU) or hardware or software control logic, read only memory (ROM), and/or other types of nonvolatile memory. Additional components of the information handling system may include one or more disk drives, one or more network ports for communicating with external devices as well as various input and output (I/O)

devices, such as a keyboard, a mouse, and a video display. The information handling system may also include one or more buses operable to transmit communications between the various hardware components.

5 Each node 101 has four processors (CPUs) 104. The number of processors is for purposes of example; a node 101 could have a single processor or some greater number of processors.

10 Each processor 104 may have the structure and function of conventional processors currently in use or of those to be developed. Input and output signals relevant to this description are shown; of course, a typical processor has many other input and output signals.

15 One output from each processor 104 is a THERMTRIP signal. A THERMTRIP signal from any processor indicates that the processor has overheated above a predetermined temperature. As explained below in connection with FIGURES 2 and 3, THERMTRIP signal from any overheated
20 processor 104 results in a reset of system 100. Upon reset, the system 100 is operational except for the node 101 associated with the overheated processor 104.

25 The THERMTRIP signal is often associated with the family of processors manufactured by Intel Corporation. However, it should be understood that any "thermal trip" signal from a processor indicating an overheating condition would be equivalent to the THERMTRIP signal.

30 A second output from each processor 104 is a PROCHOT signal. As described in the background, the PROCHOT signal may cause an affected processor 104 to reduce its

processing speed if its temperature reaches a certain level.

A THERMTRIP signal and a PROCHOT signal from each processor 104 is delivered to temperature monitor 103.

5 Temperature monitor 103 comprises logic circuitry (hardware, firmware, or instruction-based processing) that implements the functional aspects of temperature monitor 103, described below. Temperature monitor may be implemented as a programmable logic device.

10 The remaining elements of system 100 are typical of a server system. Each processor 104 is connected via a front side bus 105 to a Northbridge 106, which provides the interface to memory elements 107. A cache controller 108 handles caching operations.

15 FIGURE 3 illustrates temperature monitor 103 and its interconnections. Nodes 101 are the same as those illustrated in FIGURE 1, each node 101 having four processors 104. The THERMTRIP and PROCHOT signal connections between processors 104 and temperature
20 monitor 103 are direct wired connections.

Each node 101 has an associated voltage control module 21, connected between a power supply (not shown) and the power input to the processor 104. In the example of this description, voltage control modules 21 are
25 referred to as voltage regulator modules (VRM A and VRM B), but any voltage control circuitry capable of receiving an enable signal to control the voltage supplied to processors 104 is adequate for purposes of the invention. Like conventional voltage regulator
30 modules, each module 21 is operable to regulate the

voltage supplied to the processors 104 of its associated node 101 (Node A or Node B).

An enable signal is delivered from temperature monitor 103 to each voltage control module 21, and
5 determines whether or not the module 21 delivers voltage to its processors.

Temperature monitor 103 also delivers a system power signal to system control chipset 23. This system power signal permits temperature monitor 103 to report any
10 power shut down (such as a shut down resulting from a THERMTRIP signal) to chipset 23.

Chipset 23 may be the same as Northbridge 106 of FIGURE 1, but may also be whatever "system control unit" system 100 uses to generate a reset signal. In addition
15 to generating a reset signal, chipset 23 may use the report from monitor 103 in any additional desired manner, such as by displaying or otherwise communicating the shut down and data about the shutdown (such as date, time, and processor identification) to an operator. Chipset 23 may
20 also have any of the other functions associated with chipsets typical of server systems.

FIGURE 3 illustrates a method of using a THERMTRIP signal during run time of a multi-node server system 100, when one or more of its processors 104 overheats and
25 asserts a THERMTRIP signal. Steps 31 - 33 of the method are implemented by the logic circuitry of temperature monitor 103. Step 34 is performed by the chipset 34, triggered by the system power signal delivered from temperature monitor 103.

In Step 31, temperature monitor 103 receives the THERMTRIP signal from the overheated processor 104. In Step 32, temperature monitor 103 responds to a THERMOTRIP signal by turning off the enable signal to the voltage
5 control module 21 associated with the node 101 of the overheated processor 104. The enable signal remains in this off state regardless of the automatic resetting in Step 34.

In Step 33, temperature monitor 103 reports the
10 overheated event to chipset 23, using the system power signal. This report triggers a reset signal from chipset 23 to all processors 104. The reporting signal may include an identification of which node and/or processor 104 delivered the THERMTRIP signal, and may further
15 include data such as the date, time, and temperature during the processor failure.

In Step 34, chipset 23 responds to the report by delivering a reset signal to processors 104. As a result of the reset signal, all processors 104 are restarted in
20 the node 101 that did not contain the overheated processor. Because its power is not enabled, the node 101 with the overheated processor remains shut down until manually restarted by a technician or other operator.

Although the disclosed embodiments have been
25 described in detail, it should be understood that various changes, substitutions and alterations can be made to the embodiments without departing from their spirit and scope.